SEER*DMS Auto-Consolidation and Validation Work Group
Teleconference Summary
January 4, 2022
1:00 to 2:30 p.m. ET

Representatives from the NCI, IMS, the Scientific Consulting Group, Inc. (SCG), and 12 cancer registries participated in the SEER*DMS Auto-Consolidation and Validation Work Group (WG) conference call on January 4, 2022. Participants included:

REGISTRIES:

California Central
Connecticut
Detroit
Illinois
Iowa (Bobbi Matt, WG co-chair)
Kentucky
Louisiana
Minnesota
New Jersey
Seattle
Utah

NCI: Peggy Adamo, Lois Dickie, Marina Matatova

IMS: Suzanne Adams, Linda Coyle, Nicki Schussler

SCG: Carolyn Fisher, rapporteur

Action Items

Participants agreed to the following action items:

- Linda will create a Squish issue to evaluate the new lymphovascular invasion (LVI) default code for the borderline Benign Brain/central nervous system (CNS) auto-consolidation rule and then will recommend that registries run the task across their databases.
- IMS will create data searches for Prostate-Specific Antigen (PSA) Lab Value to identify cases where manually consolidated values differ from what the auto-consolidation logic would select. IMS will notify registries when ready to review. (See Tech Support #10344).
- IMS will develop a data search for Extent of Disease (EOD) Regional Nodes for the Liver schema showing cases where the CTC is different from what the proposed logic would have picked. (See Tech Support #10343).
- IMS and the registries will review the sentinel LN auto-consolidation logic for breast and melanoma cases diagnosed in 2018 and beyond. (See Tech Support Issue #10354).

IMS Updates

LVI Default Code

Linda noted that IMS has made changes based on discussions in the WG and issues submitted via Squish. A recent change was changing the LVI default code for borderline Benign Brain/CNS to 8 for 2010 cases forward. Registries should run the auto-consolidation rules on the test server before the SEER*DMS database as a precaution. Linda will create a Squish for testing the new rule.

Known Over Unknown Rules: Metastasis (Mets) at Diagnosis (DX)

IMS efforts have focused on implementing known over unknown rules for every field that is set when auto-building a CTC from a NAACCR abstract (or abstract). The new rules will allow registry staff to build a CTC from a pathology report and then link the abstract with the same results as building the CTC from an abstract and then link to a pathology report. IMS recently deployed the Mets at DX fields.

- Currently, if a CTC is built from a path report, and an abstract is linked later, many conflicts are generated which need manual consolidation. Because pathology reports are received quicker than abstracts (within days of the Date of Diagnosis), IMS is seeking a way to reduce manual reviews when building from a path report. The first step in this process is to implement all known over unknown rules in any field set when an abstract is consolidated into a CTC is built.

- IMS is working to implement the logic in Autobuild for pathology reports and abstracts.

- This process will be implemented first at the Louisiana registry (using the test server) from January to February 2022. A Squish issue will be made available to update other registries on the progress of this test.

- Some registry-specific fields might be needed.

SEER and National Cancer Database (NCDB) COVID-19 Auto-Consolidation Rules

Two sets of COVID fields exist: SEER COVID-19 and NCDB COVID-19. Registries were asked to collect SEER COVID-19 fields (COVID-19 Abstraction Guidance) as text (unstructured data) because the data items could not be added to the NAACCR standard quickly when collection started in 2020. Coded data items were added to abstracting tools, such as SEER*Abs, in registries that maintain the abstracting tool used by their abstractors.

- Polishers have been in place for some time to convert the record text into the coded fields. IMS has been working on a consolidation rule to place the COVID-19 record values into the combined Patient Set. In December 2021, IMS completed implementation of these auto-consolidation rules to incorporate these values into the Patient Set.

- Linda is implementing the new rules across old data sets within each registry, setting a value for blank fields and reviewing whenever a non-blank value is modified. She has posted updates in Squish issue #10231 (Changes to code COVID-19 fields in more records based on text) and registries can make comments there.

Discussion

Bobbi asked whether the SEER COVID-19 fields are still being collected by SEER and whether the National Program of Cancer Registries (NPCR) is still collecting these data. SEER continues to collect COVID-19 fields but not in the NPCR.

Linda clarified that polishers were developed and implemented over the past months to parse the COVID-19 text and set values.

Mona Highsmith (Minnesota registry), whose registry is continuing to collect COVID-19 data beyond 2020, asked whether any SEER registries had completed linkages with their state's infectious disease

databases and, if so, how these data were imported into SEER*DMS. Minnesota has submitted a grant to study the long-term effects of COVID-19, and linkages with infectious disease databases are planned. Bobbi noted that Iowa had originally asked for linkage with infectious disease data at the beginning of the COVID-19 pandemic, but that was not an option then. Mei-Chin Hsieh (Louisiana registry) noted that her registry's COVID-19 data has been linked to the state's infectious disease data, including information on COVID-19 cases identified through a PCR test that was positive for the SARS-CoV-2 antibody. The registry performs mass updates using NAACCR abstracts. Linda added that IMS could help with making sure the values obtained by linkage are not overwritten during auto-consolidation.

Heather Stabinsky (New Jersey registry) sought clarity on SEER requirements regarding the COVID-19 SARS-CoV-2 fields. Linda responded that the COVID-19 abstracting rules were posted on the SEER website in 2020. Registries were asked to enter a specific format for capturing COVID-19 information, which was incorporated SEER*DMS using a polisher.

SSDIs: PSA Lab Value Auto-Consolidation Logic

Suzanne discussed the first SSDI auto-consolidation logic for the proposed PSA laboratory value (Lab Value). A modified version of this logic could be applied to similar SSDIs. The following guidance from the NAACCR *SSDI Manual* need to be addressed in developing this coding logic: (1) record the last pre-diagnosis PSA lab value prior to diagnostic biopsy of prostate and treatment; (2) record the adjusted PSA value in the medical record, if available, over the lab value; (3) PSA is a prognostic factor required for American Joint Committee on Cancer (AJCC) staging and affects the stage group in most cases; and (4) a known lab value takes priority over codes XXX.2 and XXX.3.

The dates for the PSA values are unknown, thus auto-consolidation based on time relative to diagnostic biopsy and treatment will not be possible. Adjusted PSA values are rarely documented in the medical record (according to the registries), and it is unknown whether the adjusted lab value would be higher or lower. Therefore, adjusted PSA lab values will not be taken into consideration when auto-consolidating.

After reviewing the *SSDI Manual*, the WG Administrative team proposed the following prioritization of codes from highest to lowest:

1. 0.1–999.9, XXX.1. If there are conflicting values within this group, act based on ranges 0.1– 9.9; 10.0–19.9 or 20.0–XXX.1.
    a. If known values are in the same range, take the higher value.
    b. If known values are in different ranges, perform manual review.
2. XXX.3
3. XXX.2
4. XXX.7
5. XXX.9

Discussion

Details are available in Squish #10344. Marina recommended that registries review the Squish issue and provide feedback. Linda suggested expanding the statement in Priority code 1 to "take action based on ranges that affect staging."

EOD Proposed Auto-Consolidation Logic: Liver

Suzanne presented the proposed EOD auto-consolidation logic (Squish #10343) for liver—primary tumor, nodes, and metastasis (TNM)—derived from the *EOD Consolidation Manual* published in December 2021. The EOD liver will be the first EOD Schema to be auto consolidated and will serve as a model for future logic.

For EOD Primary Tumor, SEER*DMS registrars should set the CTC value by:

1. Selecting records with a known and valid Schema ID [#3800] and known values for EOD Primary Tumor.
2. Assigning the selected records into priority groups (1–4) based on Commission on Cancer (CoC) Accredited Flag (#2152) and Surgery of Primary Site (#1290). Refer to Squish 10343 for further details.
3. Processing the priority 1 group first. If a value is not set, apply the logic to each group in order of priority (refer to Squish 10343 for further details).
4. Setting the CTC value to 999 if no value can be set.

Discussion

Bobbi asked whether Step 2 priority groups 2 (CoC-accredited flag) and 3 (surgery/primary site) should be reversed because a CoC non-accredited hospital performing surgery would have better information. Other participants agreed, so the priority groups were revised.

Tiffany Janes (Seattle registry) suggested including criteria in Step 1 indicating that the Schema IDs should match the CTC. In addition, participants suggested including criteria to perform a manual review if Schema IDs do not match another record/CTC and have known EOD primary tumor. Suzanne noted the changes and will update the Step 1.

For EOD Nodes, SEER*DMS registrars should set the CTC value by:

1. Select records with a known and valid Schema ID [#3800] and known values of EOD Nodes.
2. Assign the selected records into priority groups (1 or 2) based on CoC Accredited Flag (#2152).
3. Process the priority 1 group first. If a value is not set, apply the logic to each group in order of priority (refer to Squish 10343 for further details).
4. If no value could be set, then set CTC value to 999.

Discussion

Suzanne confirmed that all EOD values need not originate in the same record because information can be obtained from different facilities performing the surgeries. Participants agreed that the prioritization of records should also be based on Scope of Regional Lymph Node (LN) Surgery.

Angela Veach (Illinois registry) asked whether EOD LNs included clinical information or just pathology data. Bobbi clarified that pathology data would be the priority, noting that this field also can contain clinical data. Suzanne added that the general instructions in the *EOD Manual* address this topic.

Participants suggested incorporating neoadjuvant therapy for all data items, considering a Neoadjuvant field for 2021 data and beyond, and anticipating situations (although rare) of clinically positive LNs regardless of the surgical status.

Bobbi asked about generating two separate data searches, one with Scope of Regional LN Surgery and a second without. The results will determine the logic. Linda agreed that IMS could develop these data searches.

For EOD Mets, SEER*DMS registrars should set the CTC value by:

1. Select records with a known and valid Schema ID [#3800] and known values for EOD Mets.
2. Assign the selected records into priority groups based on CoC Accredited Flag [#2152].
3. Process the priority 1 group first. If a value is not set, apply the logic to each group in order of priority (refer to Squish 10343 for further details).
4. If no value could be set, then set CTC value to 00.

Discussion

Suzanne clarified that EOD Mets does not correlate to the DX Mets fields and are not linked. She suggested reconciling any disagreement between these two fields using edits if needed.

Sentinel LNs Positive and Examined Pairs: Priority for Auto-Consolidation

Bobbi briefly discussed the proposed logic for examining Sentinel LNs Positive and Examined pairs for breast and melanoma cases diagnosed in 2018 and beyond. Valid and invalid pairs across the possible combinations have been included. Bobbi requested IMS and the registries review the logic and provide feedback.

Upcoming SEER*DMS Meetings

The next Auto-Consolidation and Validation WG call is scheduled for April 5, 2022.